ELSEVIER

Artificial Intelligence 105 (1998) 345–354

Artificial Intelligence

Research Note

The vulnerability of the transferable belief model to Dutch books

Paul Snow [1]

P.O. Box 6134, Concord, NH 03303-6134, USA

Received 1 April 1998; received in revised form 3 July 1998

Abstract

Smets and Kennes have claimed that the transferable belief model, a decision and inference procedure based upon the Dempster–Shafer formalism, never exposes the believer to a kind of betting conundrum known as a "Dutch book". A Dutch book is constructed against the model in an elaboration of an example proposed by Smets and Kennes. A condition which permits this Dutch book is identified, and is shown to conflict with an intuition about reasonable belief revision which is not confined to the probabilist community. © 1998 Published by Elsevier Science B.V. All rights reserved.

Keywords: Transferable belief model; Dempster–Shafer; Bayesian decision making

1. Introduction

The transferable belief model [6] combines elements of two belief representation formalisms: Bayesian probabilities and the Dempster–Shafer calculus. Smets and Kennes present the transferable belief model as a *normative* construct, i.e., the model *offers advice* about how to represent beliefs, revise them based upon evidence, and apply them to choices among alternative wagering commitments.

Dempster–Shafer analysis assigns a nonnegative number, sometimes called a *mass*, to every distinct logical sentence in a finite domain of sentences which is closed under the usual connectives. The sum of all masses assigned is unity. A popular first step in the interpretation of these numbers involves *belief functions*, defined as

$$bel(X) = \sum_{Y \Rightarrow X} m(Y),$$

[1] Email: paulsnow@delphi.com.

PII: S0004-3702(98)00075-7

where $bel(\ )$ is the belief function, X and Y are sentences in the domain, and $m(Y)$ is the mass assigned to sentence Y. The semantics attributed to the belief functions varies among authors, but a common theme is that if $bel(X) \geqslant bel(Z)$, then X is no less credible then Z, with the precise meaning of credibility being one of the things about which authors differ.

Formally, probability distributions and Dempster–Shafer representations are closely related. Kyburg [4] has shown that the Dempster–Shafer belief functions can be interpreted as numeric bounds in a system of simultaneous linear inequality constraints

$$p(X) \geqslant bel(X) \quad \text{for all sentences } X \text{ in the domain,}$$

$$\sum_X p(X) = 1,$$

whose solutions are a set of ordinary probability distributions.

For belief revision, Bayesians use Bayes' rule while Dempster–Shafer adherents use Dempster's rule. Kyburg notes an interesting relationship between the two approaches.

Suppose some sentence E in the domain of interest were revealed to be true, a natural representation of "observing evidence" in both formalisms. One could then use Bayes' rule to revise the set of probability distributions described in the last paragraph, yielding a new set of probabilities $p(X \mid E)$. One could also use Dempster's rule to derive new belief functions $bel(X \mid E)$. Kyburg shows that for any sentence X, if

$$p(X \mid E) \in [L_{\mathrm{B}}, U_{\mathrm{B}}] \quad \text{for all } p(\ ) \text{ in the Bayesian set}$$

then

$$L_{\mathrm{B}} \leqslant bel(X \mid E) \leqslant U_{\mathrm{B}}.$$

If one then considered the probabilistic interpretation of $bel(X \mid E)$, the set of probability distributions whose typical constraint is

$$q(X) \geqslant bel(X \mid E)$$

and examined the intervals implied by those probabilities

$$q(X) \in [L_{\text{D-S}}, U_{\text{D-S}}]$$

for all $q(\ )$ in the Dempster–Shafer interpretative set where the intervals are tight, then

$$L_{\mathrm{B}} \leqslant L_{\text{D-S}} \leqslant U_{\text{D-S}} \leqslant U_{\mathrm{B}}.$$

In words, the Dempster–Shafer intervals are always at least as specific as the corresponding Bayesian intervals. It is routine for the Dempster–Shafer intervals to be strictly more specific.

The transferable belief model assumes that beliefs can be represented according to the Dempster–Shafer conventions, and that when evidence is observed, beliefs change according to Dempster's rule. For gambling purposes, the transferable belief model selects one particular probability distribution from among the solutions of the linear constraint interpretation of the Dempster–Shafer representation. This single probability distribution is used to analyze any prospective wagers in the usual probabilist expected utility fashion.

Once chosen, a probability distribution continues to regulate gambling choices until new evidence is *actually* observed. This is an important point, since wagers may depend upon whether or not some *potential* evidence is observed in the future. If new evidence were observed during the life of the wager, then this might change the observer's beliefs and so influence expectations about the availability of pay-offs which are contingent on events or revelations which will occur after the new evidence is seen. The transferable belief model includes the rule that the credibilities of all uncertain events are evaluated according to the beliefs encoded in the probability distribution which is in force when each wagering decision is made.

To explain this rule, Smets and Kennes (at 207) distinguish "factual" conditioning from "hypothetical" conditioning. In factual conditioning, if some sentence A is learned to be true, then the model performs a Dempster's rule belief revision based upon A, and chooses a probability distribution according to the result. In hypothetical conditioning, A is not known to be true, but a bet can be offered which will be "called off" (i.e., the bet's pay-off is the *status quo*) if A turns out to be false. The model does *not* perform a Dempster–Shafer belief revision based upon A in order to evaluate such bets, but rather uses a Bayesian analysis of the current chosen probability distribution.

So, for instance (as discussed in Section 4.4 of Smets and Kennes [6]), one might begin a problem episode with the knowledge that there are three candidates for a job, and have some beliefs about their relative prospects. These beliefs are represented in the Dempster–Shafer fashion, which leads to the selection of a probability distribution, and one may place bets according to its counsel. Later on, one learns for certain that one of the candidates has not been hired. This causes a Dempster's rule revision followed by the selection of a new probability distribution to which newly offered bets are referred for evaluation. Had one *not* actually learned that the candidate was eliminated, but had merely been offered a bet between the other two candidates, then the original probability distribution would continue in force to evaluate the proposed wager, and Bayes' rule would determine the relevant odds.

2. Dutch book arguments

Smets and Kennes motivate the revision conventions just described as a precaution against *Dutch books*. A Dutch book is a suite of gambles, each of which is acceptable to the believer, but which collectively commit the believer to a situation which is unambiguously inferior to that obtained by simply declining the entire suite. Dutch books are sometimes briefly characterized as agreements to place bets in such a way as to ensure a loss for certain, regardless of how the uncertain events in question turn out.

That brief description may be misleading, however. Dutch book arguments rarely refer to any plausible threat. If concern about losses of this kind were a practical problem, then perhaps one could simply resolve to decline commitments which lead to such losses, regardless of any other beliefs one might hold [1]. Unproblematic implementation of this resolution assumes, of course, that the believer appreciates the predicament in timely fashion, which may not be the case.

In any event, practical loss is not the point. What is being evaluated is the advice offered by a normative formalism. The basis of the evaluation is a conflict with common sense, a classic comparison of a shallow problem representation with a deeper one. If a wagering problem is described one way, then the formalism counsels some course of action. If the same schedule of contingent pay-offs is described in a different way, then common sense counsels the believer to do something else. It is irrelevant whether common sense prevails in the end or not; the rub is that unacceptable advice was offered at all by a formalism whose purpose is to offer good advice.

It is uncontroversial that the usual kind of Bayesian methods do not suffer from Dutch books. Although it has been conjectured from time to time that only Bayesian belief revision methods are invulnerable to Dutch books, this is not a theorem. (See [2] for a narrative review of the voluminous literature on these points.)

The Smets and Kennes position is, in brief compass, that the transferable belief model handles all wagers in the same fashion as an orthodox Bayesian would. Orthodox Bayesianism is free of Dutch books, hence the transferable belief model should be, too. No theorem on this point was proven, although the invulnerability of the model to Dutch books was asserted as true.

As it turns out, some wagers are not handled in the same fashion as an orthodox Bayesian would. A betting problem is presented which slightly extends an example offered by Smets and Kennes. A Dutch book is constructed against the transferable belief model in this problem. A feature of the model and of the probabilistic constraint interpretation of Dempster–Shafer belief functions which allows this construction is identified. This feature is shown to conflict with a principle of inductive reasoning whose admirers within the artificial intelligence community are not confined to probabilists.

3. The example problem

> Peter, Paul and Mary are candidates for a job. Exactly one of them will be hired. The award process is as follows. A fair coin is tossed. If the coin lands showing heads, then Mary gets the job. If the coin lands showing tails, then Mary does not get the job. The procedure for deciding between Peter and Paul in that case is not disclosed. The successful candidate will be revealed in the following manner. The name of one *unsuccessful* candidate will be announced promptly after the hiring decision is made. The name of the other unsuccessful candidate will be announced after a brief interval. The procedure for deciding the order of announcements is by lot, with each unsuccessful candidate having an equal chance of being named first.

The problem is similar to the Smets and Kennes "Mr. Jones Murder Mystery" in which Peter, Paul and Mary are contract killers suspected of a homicide and Peter turns out to have an alibi. The uncertainties in the hiring procedure used here exactly parallel those of the murder mystery. In the original version, however, Peter's alibi is apparently an unexpected development and so is not ripe for betting. In the present version, the "alibi" event is scheduled and any candidate could have one. The announcement procedure described is similar to the method actually used in the annual Miss America contest to disclose who has gotten that particular job.

Before the first announcement is made, Dempster–Shafer conventions assign the masses:

Mary gets the job	0.5,
Peter or Paul gets the job	0.5.

In the event that Peter's name is announced first, Dempster's rule revision yields:

Mary gets the job	0.5,
Paul gets the job	0.5.

Peter and Paul's situations are symmetric, so if Paul's name is announced first, then the belief masses become:

Mary gets the job	0.5,
Peter gets the job	0.5.

How these numbers are arrived at is fully explained in the Smets and Kennes paper, or any exposition of the Dempster–Shafer method. We can construct a Dutch book without detailing the beliefs which would prevail should Mary's name be announced.

The transferable belief model chooses the probability distribution

$$p(\text{Mary gets the job}) = 0.5,$$
$$p(\text{Peter gets the job}) = 0.25,$$
$$p(\text{Paul gets the job}) = 0.25$$

for use before any name is announced. If Peter's name is announced first, then the chosen probability distribution is:

$$p(\text{Mary gets the job}) = 0.5,$$
$$p(\text{Paul gets the job}) = 0.5.$$

If Paul's name is announced first, then the chosen distribution is:

$$p(\text{Mary gets the job}) = 0.5,$$
$$p(\text{Peter gets the job}) = 0.5.$$

Of these three distributions, only the first is peculiar to the transferable belief model. In the other two cases, the probability distributions are the only solutions of the system of simultaneous constraints built from the posterior Dempster–Shafer belief functions which was described earlier.

The objective probability that Mary's name will be called first is 0.25: the fair coin must show tails, and thereafter she must be picked in the drawing of lots. This is the only announcement probability which will figure in subsequent discussion. The value is consistent with the probability distribution chosen by the transferable belief model.

One could, if one chose, perform the Dempster–Shafer analysis by explicitly representing the prior knowledge of the announcement schedule. That is, one would at first assign the weights:

(Mary gets the job, Peter is dropped) = 0.25,

(Mary gets the job, Paul is dropped) = 0.25,

(Peter or Paul gets the job, Mary is dropped) = 0.25,

(Peter or Paul gets the job, the other man is dropped) = 0.25

all of which coincide with objective probabilities. One would then revise beliefs according to the information

(Mary or Paul gets the job, Peter is dropped) = 1,

or else

(Mary or Peter gets the job, Paul is dropped) = 1.

Applying the rules of the transferable belief model, the wagering probabilities discussed earlier all come out the same.

4. A Dutch book

The analysis presented in this section assumes that the player uses a linear utility function, that is, the desirabilities of wagers are in the same order as their average monetary pay-offs. The choice of the linear utility is conventional for small-stakes wagers, and promotes ease of exposition. An isomorphic example could be constructed using arbitrary utility units, but no gain in readability would be achieved.

The Dutch book is made as follows. Before the first name is announced, the player is offered the lottery which pays:

a gain of \$1.01 in the event that Mary gets the job,

a loss of \$1 in the event that Peter or Paul gets the job.

This lottery is strictly acceptable under the linear utility function since

(0.5 times \$1.01) > (0.5 times \$1).

Thus, the transferable belief model advises the player to accept this wager. An orthodox Bayesian would do likewise. The bet is more than fair.

That deal concluded, the bookie then offers the following proposition, still prior to the announcement of the first name.

> In return for your payment to me of \$0.11 right now, I agree that if Mary's name is announced, then I shall forgive your obligations under the wager which you have just made.

This is a generous offer. In the event that Mary's name is announced, the player would hold a debt of \$1 for certain—the player would simply have lost the original bet. The right to rid oneself of this obligation would have an expected worth under pre-announcement beliefs of

0.25 times \$1 = \$0.25

which exceeds the asking price of \$0.11. The transferable belief model offers the same advice as the Bayesian orthodoxy. The player should pay the \$0.11, and acquire the "hedge" option. It is a good buy, plain and simple. The bookie has been generous to a fault throughout the story.

The hiring procedure ensues. If Mary's name is announced, then the player ends up losing the \$0.11 which was paid for the option. If a man's name is announced, then the bookie makes one last offer:

> I shall pay you \$0.01 right now, if you surrender your rights and obligations under the wager you now hold.

This is not a "hypothetical" wager offered before a name is announced. It is a new "factual" wagering opportunity, available only if a man's name has actually been announced, and disclosed only after the announcement. The transferable belief model player owns a lottery which offers, according to the beliefs then prevailing, an even chance to gain \$1.01 or else to lose \$1. The expected value of this lottery is less than \$0.01, so the advice of the model will be to make the exchange. The player who accepts this advice sells the lottery for a gain of \$0.01, less the \$0.11 paid for the option, for a net loss of \$0.10 on the series of transactions.

It is uncertain whether the player would notice the Dutch book as it unfolded. If Mary's name is called, then the player has lost a bet in an ordinary way, and knows nothing about the other offer which will never be made. If, say, Peter is culled, then the bookie has no incentive to point out that the same option would have been proposed if Paul had been eliminated. Even a suspicious player would have no clear evidence of trickery without the bookie's testimony. ("Dutch what? No, Paul's a friend, and I do not want it to get out that I was betting against him personally. If Peter had made it this far, I would not care".)

A Bayesian avoids the Dutch book, of course. Regardless of how the Bayesian allocates probabilities between the two men originally, then it is easy to show that in at least one case where a man is eliminated, Mary's prospects will be rated no less than 2:1 against the surviving man. The high confidence of the Bayesian regarding Mary's chances in that case will make the original lottery worth more to the player than the \$0.01 exchange price, and so the terms of the lottery will be carried out. The Bayesian finishes up in that case with either a net gain of \$0.90 if Mary gets the job, or else a total loss of \$1.11 if she does not.

There is no need to allocate belief between the two men in order to participate in this favorable game and avoid the Dutch book. The original lottery requires no allocation, since the player's loss is the same for either man's success. The same is true if Mary's name is called. It is simple to verify that the objective probability that Mary gets the job given that *some* man's name is called is two thirds. One could just ignore *which* man's name has been called. The player need not believe that Mary has 2:1 odds of prevailing over whichever specific man survives; that depends on the unknown post-tails job selection procedure. The policy of "picking Mary" on the occasion that either man's name is called, however, is objectively a 2:1 commitment. Betting really can be different from believing.

But the transferable belief model player is advised to bet according to one's beliefs. The player could simply decline the model's advice if a man's name is called first. It is difficult to see why anyone committed to the model's concept of belief revision would want to do that, even if the player suspects sharp practice. If a man's name has been called, then one's beliefs are evenly split if the model is right. The \$0.11 is gone (with no basis for regret, the purchase was shrewd), the \$0.01 for certain *is* the better prospect, or so the model says. Why not take it?

Surely, the player ought not to decline the exchange in order to "avoid" a Dutch book. To succumb to a Dutch book is nothing more or less than to put oneself in situation like this. The advice has been given. What one does about it is beside the point.

5. What is learned when a rival is eliminated

The feature of the transferable belief model which is exploited in the Dutch book of the previous section is that the model does not allow the elimination of a male rival to preserve the ordinal estimate of Mary's prospects relative to the surviving man. This belief revision behavior conflicts both with Bayesianism and with an intuition that is not peculiar to probabilists.

Before any name was announced, in all probability distributions which satisfy the constraints imposed by the Dempster–Shafer belief functions,

$$p(\text{Mary gets the job}) \geqslant p(\text{Paul gets the job}),$$

$$p(\text{Mary gets the job}) \geqslant p(\text{Peter gets the job})$$

and at least one inequality is strict. In the particular probability distribution chosen by the transferable belief model, both inequalities are strict. After a male rival is eliminated, however,

$$p(\text{Mary gets the job}) = p(\text{Paul gets the job}),$$

or else

$$p(\text{Mary gets the job}) = p(\text{Peter gets the job}).$$

This subtle but fraught ordinal shift cannot happen in the Bayesian revision schema, as is well known and easily shown using familiar properties of probability. Suppose we have three sentences x, y, and z. If z is exclusive of the other two sentences, then

$$x \wedge \neg z \Leftrightarrow x \quad \text{and} \quad y \wedge \neg z \Leftrightarrow y$$

so

$$p(x) = p(x \wedge \neg z) \quad \text{and} \quad p(y) = p(y \wedge \neg z).$$

If further we have $p(z) < 1$, so $p(\neg z) > 0$, then by a widely used general rule,

$$p(x \mid \neg z) \geqslant p(y \mid \neg z) \quad \text{if and only if} \quad p(x \wedge \neg z) \geqslant p(y \wedge \neg z)$$

leading to the conclusion that if $p(z) < 1$,

$$p(x \mid \neg z) \geqslant p(y \mid \neg z) \quad \text{if and only if} \quad p(x) \geqslant p(y).$$

In words, eliminating rivals preserves the prior credibility ordering of the survivors.

The remaining case, where the probability of z is unity, is not relevant to the present discussion. The men must survive the coin flip to succeed. The case is interesting when it does come up, though, since finding that $\neg z$ obtains (in a finite domain of alternatives) would be "learning the impossible"—and that one's beliefs are simply wrong.

The notion that eliminating rivals preserves the order of the survivors can be motivated from other intuitions besides those specific to probability. A staple of the case-based reasoning literature is a principle which may be stated

> If knowing c would lead to some conclusion, and if knowing d would lead to the same conclusion, then knowing $c \vee d$ also leads to that conclusion.

Neither probabilists nor all case-based reasoners would assent to that statement in its full generality, but the intuition behind the disjunctive principle is respectable [3].

Applying this idea to the elimination of rivals situation, suppose c is exclusive of a and of b, and that sentences are completely and definitely ordered with respect to credibility. That is, two sentences are either strictly ordered with respect to one another, or else in equipoise. Knowing c to be true would make a and b equally credible: they would both be false. Thus, both senses of weak ordering are true when c is true. So, under the disjunctive principle, whatever weak ordering prevailed between a and b when c was known to be false would also prevail when $c \vee \neg c$ were known to be true, which is to say, as soon as the problem is stated or *a priori*. The converse follows easily for a complete definite ordering of sentence credibilities by a *reductio*.

Adherence to the elimination-of-a-rival principle by itself does not necessarily provide protection against Dutch books. With some attention to the specific numbers used, if the coin which decides between Mary and the men were slightly biased in favor of heads, then the transferable belief model would succumb to the same book without violating the ordinal principle.

Nevertheless, the model's differences regarding the features of intuitive belief revision are not confined to orthodox Bayesian dogma. Its disagreement with a not-specifically-probabilist intuition is sufficient to elicit anomalous gambling advice from the model.

6. Conclusions

The transferable belief model was said to offer gambling advice of the same quality as the orthodox Bayesian method according to a criterion advanced by Bayesians themselves. The purpose of the present paper is to evaluate that performance claim, and to give some explanation of what goes wrong when the assertion is revealed to be false.

The wagering anomaly presented illustrates a questionable feature of the probabilistic constraint interpretation of the underlying belief revision method. Elimination of a rival can disrupt the credibility ordering of the surviving hypotheses, contrary to the probabilist intuition presumably needed for the interpretation to be cogent. The enhanced specificity of Dempster–Shafer revision methods discussed by Kyburg [4] is not an unalloyed advantage over Bayesian approaches, but a principled disagreement about what may be inferred from evidence. Kyburg did warn that belief functions might come to grief in gambling applications when interpreted probabilistically.

Gambling difficulty in itself does not reflect poorly on the merits of the Dempster–Shafer formalism as a belief manager. Shafer, in a commentary published with [5], suggested that wagering advice need not agree with belief management teachings. This suggestion is entirely in the spirit of the transferable belief model, with its distinct credal and pignic elements. The results presented here suggest that betting and believing might profitably be separated even more thoroughly than the modelers proposed.

References

[1] F. Bacchus, H.E. Kyburg Jr., H. Thalos, Against conditionalization, Synthese 85 (1990) 475–506.
[2] C. Howson, P. Urbach, Scientific Reasoning: The Bayesian Approach, 2nd ed., Open Court, Chicago, IL, 1993, Chapter 6.
[3] S. Kraus, D. Lehman, M. Magidor, Nonmonotonic reasoning, preferential models, and cumulative logics, Artificial Intelligence 44 (1990) 167–207.
[4] H.E. Kyburg Jr., Bayesian and non-Bayesian evidential updating, Artificial Intelligence 31 (1987) 271–293.
[5] D.V. Lindley, Scoring rules and the inevitability of probability, (with commentaries), Internat. Stat. Rev. 50 (1982) 1–26.
[6] P. Smets, R. Kennes, The transferable belief model, Artificial Intelligence 66 (1994) 191–234.